Revisions to the JDL Data Fusion Model

Alan N. Steinberg
Technical Director,
Data Fusion
ERIM International, Inc.
1101 Wilson Blvd
Arlington, VA 22209
(703)528-5250 x4109
steinberg@erim-int.com

Christopher L. Bowman
Data Fusion and Neural Networks
1643 Hemlock Way
Broomfield, CO 80020
(303)469-9828
cbowman@indra.com

Franklin E. White
Director,
Program Development
SPAWAR Systems Cente
San Diego, CA 92152
Chair, Data Fusion Group
(619) 553-4036
whitefe@spawar.navy.mi

Presented at the AIAA Missile Sciences Conference, Naval Postgraduate School, Monterey, CA
17-19 November 1998

1. Abstract

The Data Fusion Model maintained by the JDL Data Fusion Group is the most widely-used method for categorizing data fusion-related functions. This paper discusses the current effort to revise and expand this model to facilitate the cost-effective development, acquisition, integration and operation of multi-sensor/multi-source systems.

Data fusion involves combining information – in the broadest sense – to estimate or predict the state of some aspect of the universe. These may be represented in terms of attributive and relational states. If the job is to estimate the state of a people (or any other sentient beings), it can be useful to include consideration of *informational* and *perceptual* states in addition to the *physical* state.

Developing cost-effective multi-source information systems requires a standard method for specifying data fusion processing and control functions, interfaces, and associated data bases. The lack of common engineering standards for data fusion systems has been a major impediment to integration and re-use of available technology. There is a general lack of standardized — or even well-documented — performance evaluation, system engineering methodologies, architecture paradigms, or multi-spectral models of targets and collection systems. In short, current developments do not lend themselves to objective evaluation, comparison or re-use.

This paper reports on proposed revisions and expansions of the JDL Data Fusion model to remedy some of these deficiencies. This involves broadening the functional model and related taxonomy beyond the original military focus, and integrating the Data Fusion Tree Architecture model for system description, design and development.

2. What is Data Fusion? What isn't?

Data fusion is an increasingly important element of diverse weapon, intelligence, and commercial systems. Data fusion uses overlapping information to determine relationships among data (the data association function). It uses the synergistic differences in the data to improve the estimate/assessment of a

reported environment (state estimation function). As such, data fusion can enable improved estimation of situations and, therefore, improved responses to situations (Figure 1).

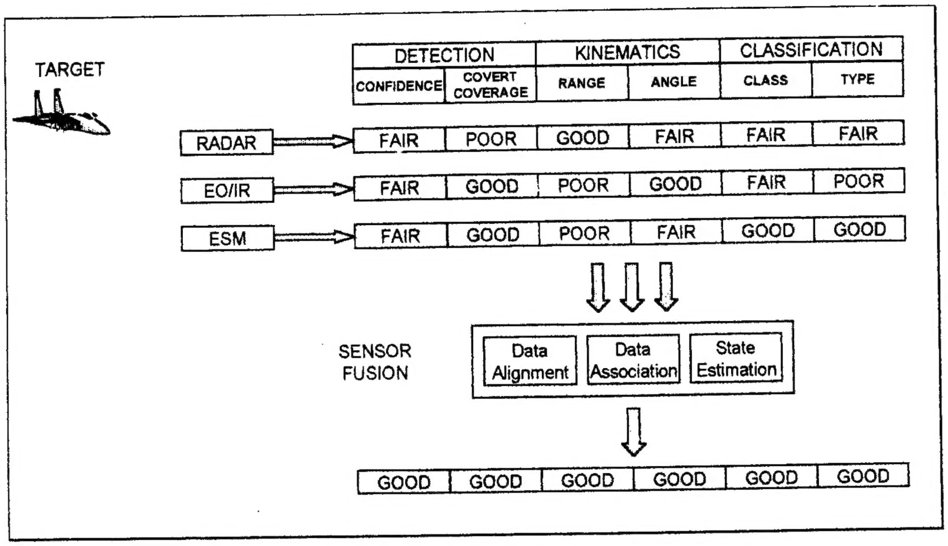

Figure 1 Data Association uses overlapping sensor capabilities so that State Estimation can exploit their complementary capabilities

Automated data fusion processes are generally employed to support human decision-making by refining and reducing the quantity of information that system operators need to examine to achieve timely, robust, and relevant assessments and projections of the situation.

Unfortunately, data fusion is a victim of its own popularity: the pervasiveness of data fusion functions has engendered a profusion of overlapping research and development in many applications. A welter of confusing terminology (Figure 2) obscures the fact that the same ground has been plowed repeatedly.

The initial Data Fusion Lexicon, produced by the JDL Data Fusion Subgroup in 1987, defined *data fusion* as

> *a process dealing with the association, correlation, and combination of data and information from single and multiple sources to achieve refined position and identity estimates, and complete and timely assessments of situations and threats, and their significance. The process is characterized by continuous refinements of its estimates and assessments, and the evaluation of the need for additional sources, or modification of the process itself, to achieve improved results* [1].

As theory and applications have evolved over the years, it has become clear that this initial definition is rather too restrictive. We need to capture the fact that similar underlying problems of data association and combination occur in a very wide range of engineering, analysis and cognitive situations. It is in this spirit that we critique the initial definition:

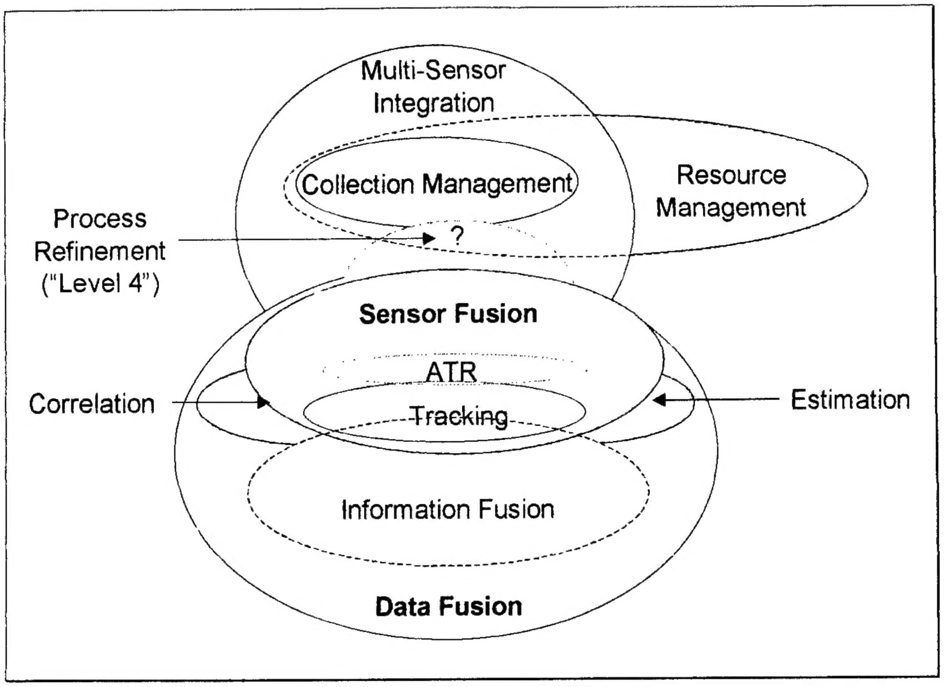

Figure 2 (Con)fusion of terminology

(a) to say that *data fusion is a process dealing with* ... suggests that there may be others. The way in which data fusion *deals with* these topics needs to be clarified;

(b) although the concept *combination of data* encompasses the broad range of problems of interest,[1] *correlation* does not. Statistical correlation is merely one method for generating and evaluating hypothesized associations among data;

(c) indeed, association is not an essential ingredient in combining multiple pieces of data. The very important recent work in Random Set models of data fusion [2,3,4] and related work in unified data fusion [5] provide generalizations that allow state estimation of multiple targets without explicit report-to-target association;

(d) since *single or multiple sources* is comprehensive, it is superfluous in a definition;

(e) the reference to *position and identity estimates* should be broadened to cover all varieties of state estimation;

(f) *complete* assessments are not required in all applications; *timely*, being application-relative, is superfluous;

(g) *threat assessment*, of course, only has application in situations where threat is a factor. We need to broaden this notion to include any assessment of the cost or utility implications of estimated situations. In general, data fusion involves refining and predicting the states of entities, aggregates of entities and their relation to one's own plans and

[1] For lack of a more general term to encompass all, we may bundle *information*, *knowledge*, etc., as subsets of *data*.

goals. These estimates can include utility or other cost estimation (e.g. the probability of survival given an estimated threat situation).

(h) Not every process of combining information involves collection management or process refinement. Thus the definition's second sentence is illustrative, not definitional.

In summary, the following concise definition is proposed:

> *Data fusion is the process of combining data to refine state estimates and predictions.*

It is fairly pointless to argue whether the term *data fusion* or some other term (e.g. one of those included in Figure 2) is an appropriate label for this very broad concept. There is no body of common and accepted usage to which we can appeal for such specialized terms. What is important is the recognition that this broad concept is an important topic for a unified theoretical approach, and therefore deserving of its own label..

Recognizing the common elements across the diversity of data combination problems can provide an enormous opportunity for synergistic development. These range from signal sorting, target tracking, multiple sensor Automatic Target Recognition and Combat Identification, Battlefield Situation Awareness (in military applications), to system fault diagnosis, medical diagnosis, criminal investigation, and economic, geopolitical and weather forecasting.

3. Data Fusion "Levels"

Of the many possible ways of differentiating among types of data fusion processes, that of the Joint Directors of Laboratories' Data Fusion Sub-Panel (now the Data Fusion Group has gained the greatest popularity. The JDL distinction among fusion "levels" (depicted in Figure 3) provides an often useful distinction among data fusion processes that relate to the refinement of "objects", "situations," "threats" and "processes."[6]

There are, however, the following concerns with the ways in which these JDL Data Fusion levels have been used in practice:

- The JDL levels have frequently been interpreted as a canonical guide for partitioning functionality within a system: do level 1 fusion first, then levels 2, 3 and 4;
- The original — and to some extent the current — JDL titles for these levels appear to be focused on tactical targeting applications, so that the extension of the concepts (e.g. of "threat refinement") to other applications is not obvious;
- For this and other reasons, the literature is rife with diverse interpretations of these levels. The levels have been taken as distinguishing any of the following: (a) the kinds of association and/or characterization processing involved; (b) the kinds of entities being characterized; (c) the degree to which the data used in the characterization already has been processed.

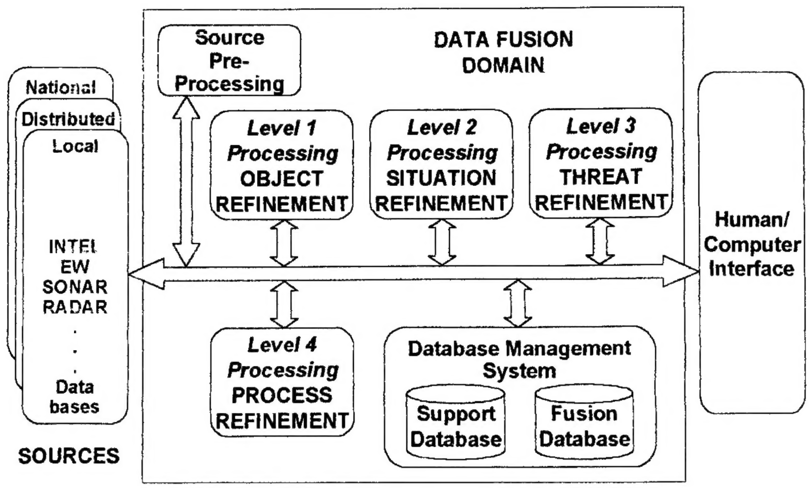

Figure 3 The JDL data fusion model (1992 version)

Let us try refining definitions for the "levels". Our objectives are to (a) provide a useful categorization representing logically different types of problems, which are generally (though not necessarily) solved by different techniques; and (b) maintain a degree of consistency with the mainstream of technical usage.

Our proposed definitions are as follows:

- Level 0 – Sub-Object Data Assessment: estimation and prediction of signal/object observable states on the basis of pixel/signal level data association and characterization;
- Level 1 – Object Assessment: estimation and prediction of entity states on the basis of observation-to-track association, continuous state estimation (e.g. kinematics) and discrete state estimation (e.g. target type and ID);
- Level 2 – Situation Assessment: estimation and prediction of relations among entities, to include force structure and cross force relations, communications and perceptual influences, physical context, etc.;
- Level 3 – Impact Assessment: estimation and prediction of effects on situations of planned or estimated/predicted actions by the participants; to include interactions between action plans of multiple players (e.g. assessing susceptibilities and vulnerabilities to estimated/predicted threat actions given one's own planned actions);
- Level 4 – Process Refinement (an element of Resource Management): adaptive data acquisition and processing to support mission objectives.

Table 1 gives a general characterization of these concepts. Note that we differentiate the levels first on the basis of types of estimation process, which typically relates to the type of entity for which state is estimated.

If the process involves explicit association in performing state estimates (usually, but not necessarily the case), there is a corresponding distinction among types of association process.

Figure 4 depicts the sorts of assignment matrices typically formed in each of these processing levels. The examples are of two-dimensional matrices, as in associating current reports to tracks.

Table 1 Characterization of the revised data fusion levels

Data Fusion Level	Association Process	Estimation Process	Entity Estimated
L.0 – Sub-Object Assessment	Assignment	Detection	Signal
L.1 – Object Assessment		Attribution	Physical Object
L.2 – Situation Assessment	Aggregation	Relation	Aggregation (Situation)
L.3 – Impact Assessment		Plan Interaction	Effect (Situation|Plans)
L.4 – Process Refinement	Planning	(Control)	(Action)[2]

Level 0 assignment involves hypothesizing the presence of a signal (i.e. of a common source of sensed energy) and estimating its state. Level 0 assignments include (a) signal detection on the basis of integration of a time-series of data (e.g. the output of an A/D converter) and (b) feature extraction from a region in imagery. In this case, a region may correspond to a cluster of closely spaced objects or to part of an object.

Level 1 assignments involve associating reports (or tracks from prior fusion nodes in a processing sequence) into association hypotheses; for which we use the convenient shorthand, 'tracks'. Each such track represents the hypothesis that the given set of reports is the total set of reports available to the system referencing some individual entity.

Level 2 assignment involves associating tracks (i.e. hypothesized entities) into aggregations. The state of the aggregate is represented as a network of relations among its elements. We admit any variety of relations to be considered – physical, organizational, informational, perceptual – as appropriate to the given system's mission. As the class of relationships estimated and the numbers of interrelated entities broaden, we tend to use the term 'situation' for an aggregate object of estimation. A model for such development is presented in [7].

---

[2] Process Refinement does not involve estimation, but rather control. Therefore, its product is a control sequence, which -- by the duality of estimation and control -- relates to a controlled entity's actions as an estimate relates to an actual state.

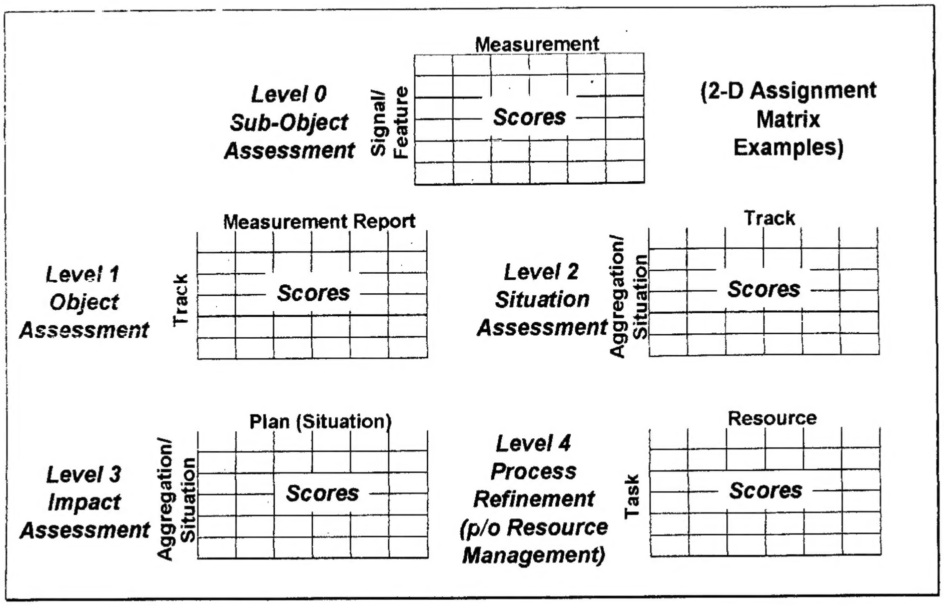

Figure 4 Assignment matrices for various data fusion "levels"

Level 3 assignment is usually implemented as a prediction function, drawing particular kinds of inferences from Level 2 associations. Level 3 fusion estimates the "impact" of an assessed situation; i.e. the outcome of various plans as they interact with one another and with the environment. The impact estimate can include likelihood and cost/utility measures associated with potential outcomes of a player's planned actions.[3]

Level 4 processing involves planning and control, not estimation. As discussed in [8], as there is a formal duality between estimation and control, there is a similar duality between association and *planning*. Therefore, level 4 assignment involves assigning tasks to resources.

An example of the relationships represented in fusion levels 0-3 is given in Figure 5. More and more, we are seeing the value of estimating entity states on the basis of context. A system that integrates data association and estimation processes of all "levels" will permit entities to be understood as parts of complex situations. A relational analysis, as illustrated in Figure 6, permits evidence applicable resolving to a local estimation problem to be propagated through a complex relational network.

Figure 7 depicts typical information flow across the data fusion "levels." Level 0 functions combine measurements to generate estimates of signals or features.

At level 1, signal/feature reports are combined to estimate the states of objects. These are combined, in turn, at

[3] Because we have defined level 2 so broadly, level 3 is actually a subset of level 2. Whereas level 2 involves estimating/ predicting all types of relational states, level 3 involves predicting some or all of the relationships between a player and his environment, to include interaction with other players' actions, given the player's action plan and that of every other player.

level 2 to estimate situations (i.e. estimations of states of aggregations of entities). It is seen that level 3 is, according to this logical relationship, out of numerical sequence. It is a "higher" function than the planning function of level 4.

Indeed, Process Refinement (level 4) processes can interact with "classical" association/ estimation data fusion processes in any of a variety of ways, managing the operation of individual fusion nodes or that of larger ensembles of such nodes, as illustrated in Figure 9 below.

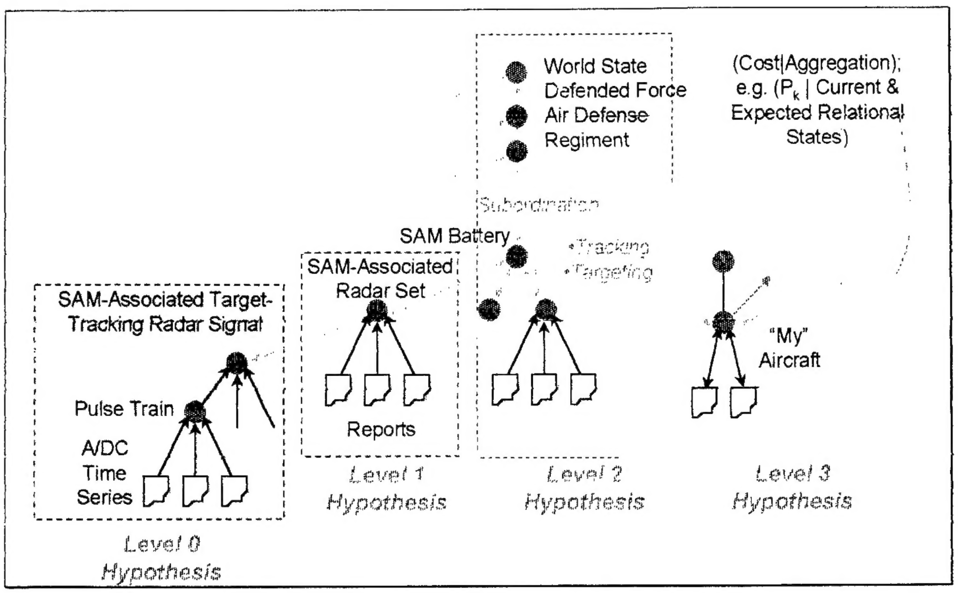

Figure 5 Multi-level inferencing example

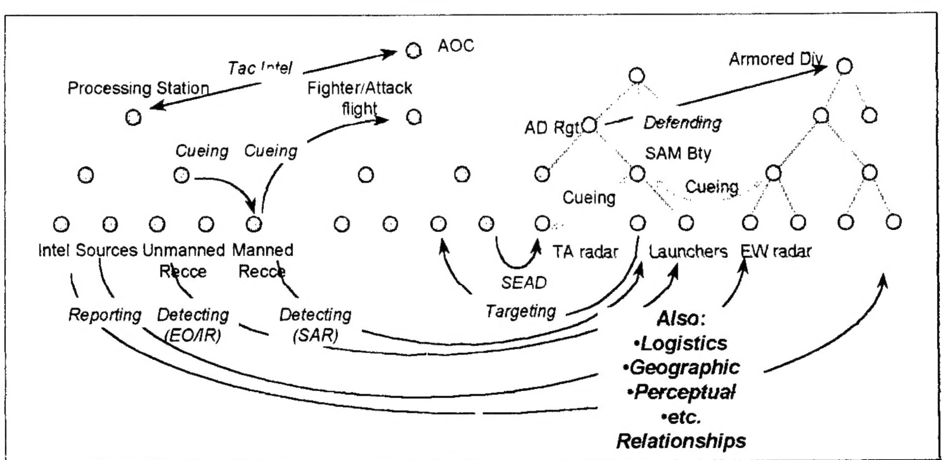

Figure 6 Understanding complex situations requires propagating evidence through complex relational networks

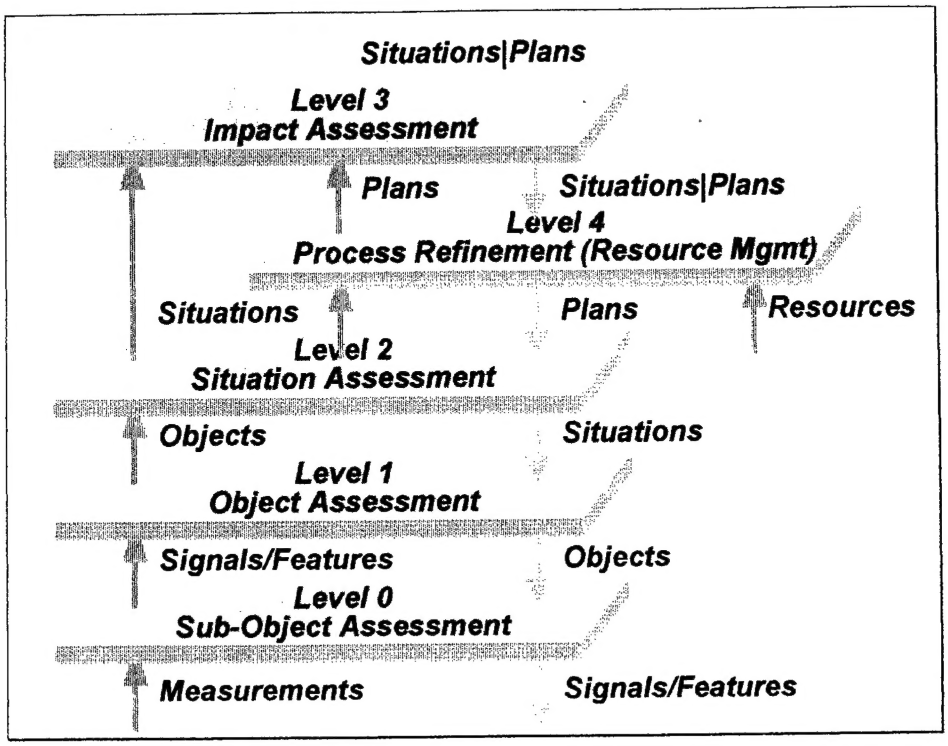

Figure 7 Logical flow among the "levels"

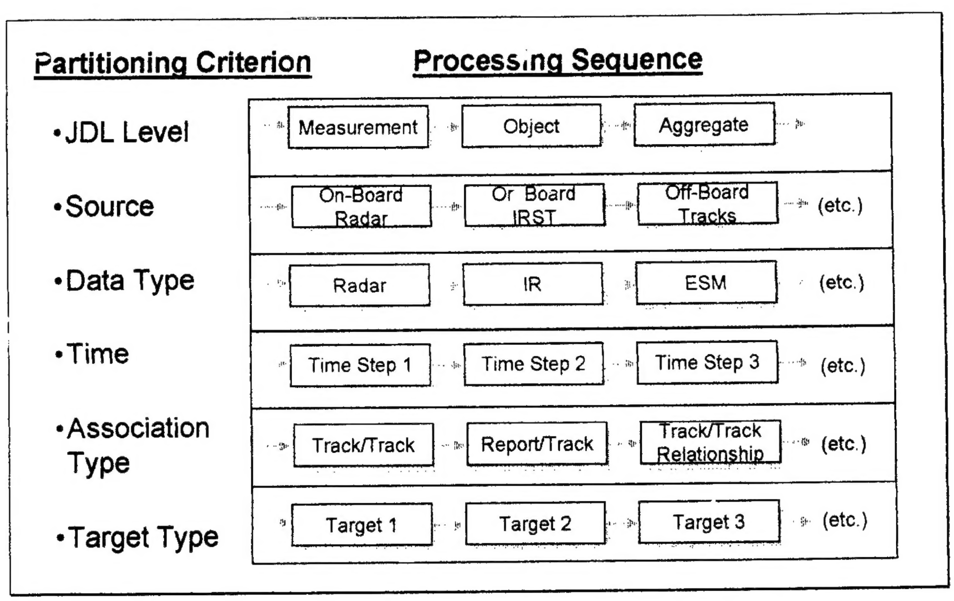

Figure 8 Alternative data fusion partitioning schemes

This point reinforces an important caveat, one too often ignored by designers of data fusion system: the Data Fusion levels are intended only as a convenient categorization of data fusion functions. They were never intended to be, nor should they be taken as a prescription for designing systems: do level 0 fusion first, then level 1, then level 2, etc. Processing should be partitioned in terms of the individual system requirements. As shown in

Figure 8, the principles of assignment and aggregation -- i.e. the principles distinguishing of levels 0-3 – do not exhaust the ways of partitioning data association/state estimation problems. Often hybrid or adaptive approaches are appropriate.

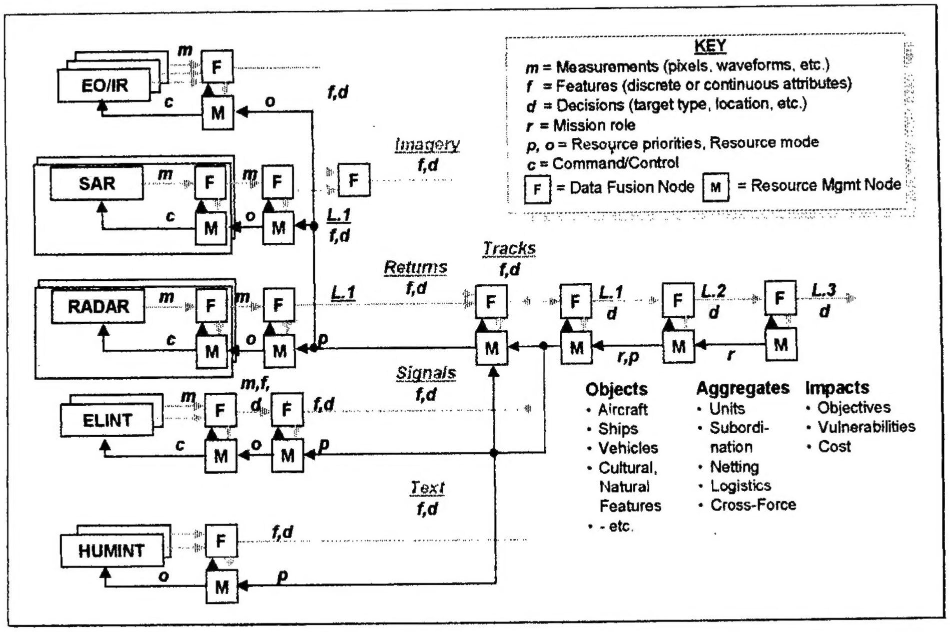

Figure 9 Integrated data fusion/resource management trees

Furthermore, diverse system requirements can drive the designer to many different solutions for integrating data fusion and resource management (to include process refinement, our so-called data fusion level 4).

Figure 9 shows the sort of highly integrated fusion/management systems appropriate to applications rapid response as part of a multi-faceted, spatially distributed, sensor/response system.[4] Such solutions are facilitated by the formal duality between data fusion and resource management, resulting in the analogous processing node paradigms for the two functions, shown in Figure 10.

[4] As one moves to the right of interlaced Fusion/Management trees as depicted in Figure 8, the Fusion/Management node pairs generally operate with broader perspectives and slower response times.

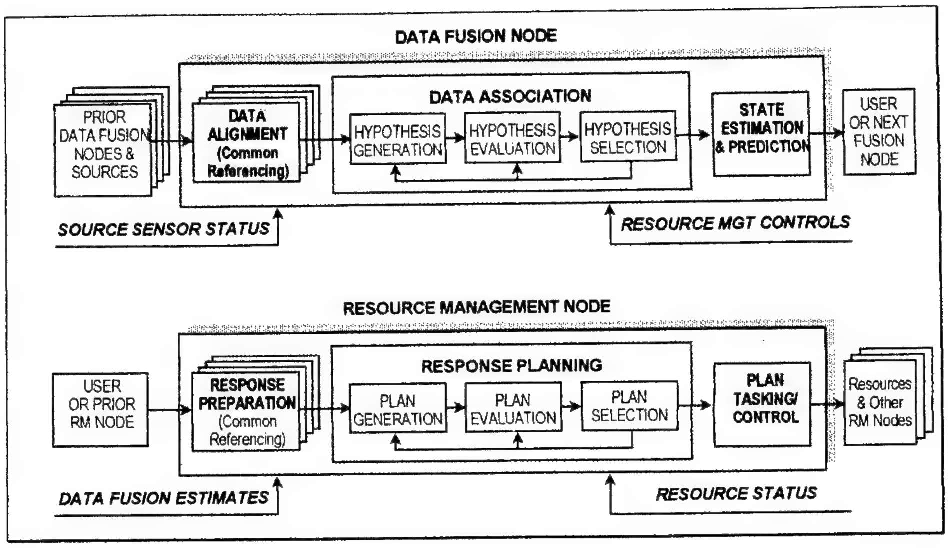

Figure 10 Duality of data fusion and resource management nodes

4. Beyond the Physical

In general, then, the job of data fusion is that of estimating the state of some aspect of the world. When that aspect includes people (or any other information systems, for that matter), it can be useful to include consideration of states in addition to the physical concerns of *who, what,* and *where*. In estimating and predicting the state of an observed target, it is common to postulate internal states (e.g. a hidden Markov process when the states are discrete). When the target is a person, or a group of people, we may think in terms of the target's informational and perceptual states as well as physical states. By *informational* state we mean the data available to the target. By *perceptual* state we mean the targets own estimate of the world state.

A person or other information system (represented by the box at the left of Figure 11) senses physical stimuli as a function of his physical state in relation to that of the stimulating physical world. These include both stimuli originating outside the person's body and those originating from within. He can combine multiple sensory reports to develop and refine what we may call "tracks" (perceived entities), perceived aggregations, and estimated/predicted interactions (levels 1-3 fusion). This ensemble of perceived entities and their interrelationships may be considered a part of the person's *Perceptual State*. As depicted in the figure, his perceptual state can include an estimation of physical, informational and perceptual states and relations of things in the world.

The person's perceptions can be encoded symbolically for manipulation, communication or storage. The set of symbolic representations available to the

person may be termed his *Informational State*.[5]

The person acts in response to his perceptual state, thereby affecting his and the rest of the world's physical state. His actions may include comparing and combining various representations of reality: his network of perceived entities and relationships. He may search his memory or seek more information from the outside." These are processes associated with data fusion level 4.

Other responses can include encoding perceptions in symbols, for storing or communicating perceptions. These can be incorporated in the person's physical actions. These, in turn, are potential stimuli to people (including the stimulator himself) and other entities in the physical world, depicted at the right of the figure.

---

[5] *Informational State* may be considered to encompass available data stores: databases, documents, etc. The notion of *Information State* is probably more applicable to a closed system (e.g. a non-networked computer) than to a person, for whom the availability of information is generally a matter of degree. The tripartite view of reality is developed by E. Waltz [8] with reminiscences of the philosopher Karl Popper. The status of "Information" as a separable aspect of reality is certainly subject to discussion. Symbols can have both a physical and a perceptual aspect: they can be expressed by physical marks or sounds, but their interpretation (i.e. recognizing them orthographically as well as semantically) is a matter of perception:

As seen in this example, symbol recognition (e.g. reading) is clearly a perceptual process. It is a form of context-sensitive model-based processing. The converse process, that of representing perceptions symbolically for purpose of recording or communicating them, produces a physical product – text, sounds, etc. Such physical products need to be interpreted as symbols before their informational content can be accessed. Whether there is anything to information in addition to these physical and perceptual aspects is not totally clear. Nor is the distinction between information and perception that between what a person *knows* and what he *thinks* (cf. Plato's *Theatetus*, in which knowledge is shown to involve true opinion plus some sense of understanding). Nonetheless, the notion of *Informational State* is useful as a topic for estimation, since knowing what information is available to an entity (e.g. an enemy commander's sources of information) is an important element in predicting his perceptual state and, therefore, in predicting changes in his physical state.

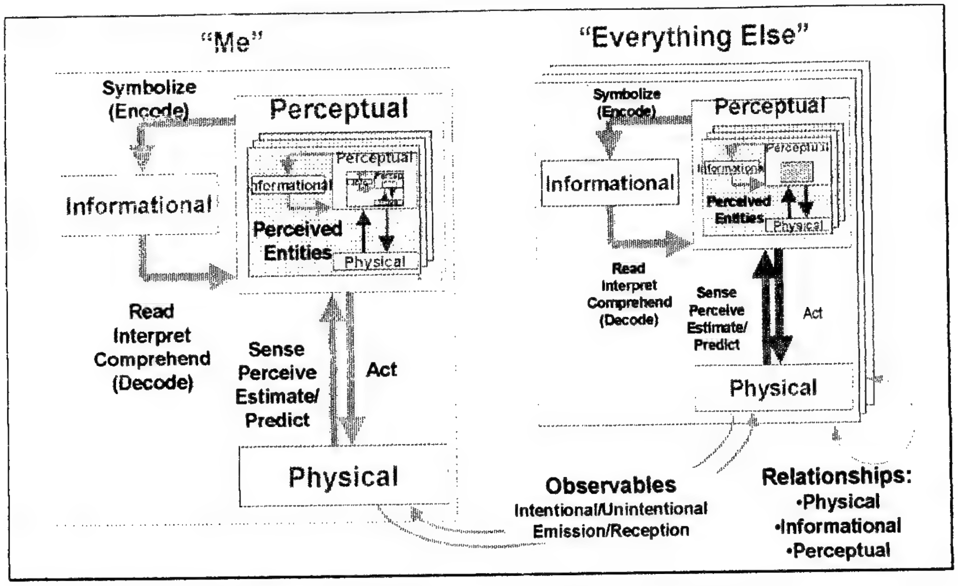

Figure 11 Entity states: three aspects

The elements of state estimation in each of theses three aspects are given in Table 2. Note the recursive reference in the lower right cell.

Figure 12 illustrates this recursive character of perception. Each decision-maker interacts with every other one on the basis of an estimate of current, past and future states. These include not only estimates of who is doing what, where and when in the physical world; but what is their informational state and what is their perceptual state (including, what do they think of *me*?).

Table 2 Elements of state estimation

Object Aspect	Attributive State		Relational State	
	Discrete	Continuous	Discrete	Continuous
Physical	• Type, ID • Activity State	• Location/ Kinematics • Waveform Parameters	• Causal Relation Type • Role Allocation	• Spatio/ Temporal Relationships
Informational	• Available Data Types • Available Data Records and Quantities	• Available Data Values • Accuracies • Uncertainties	• Informational Relation Type • Info Source/ Recipient Role Allocation	• Source Data - Quality - Quantity - Timeliness • Output QQT

Perceptual	• Goals • Priorities	• Cost Assignments • Confidence • Plans/ Schedules	• Influence Relation Type • Influence Source/ Recipient Role Allocation	• Source Confidence • World State Estimates (per *Table 2*)

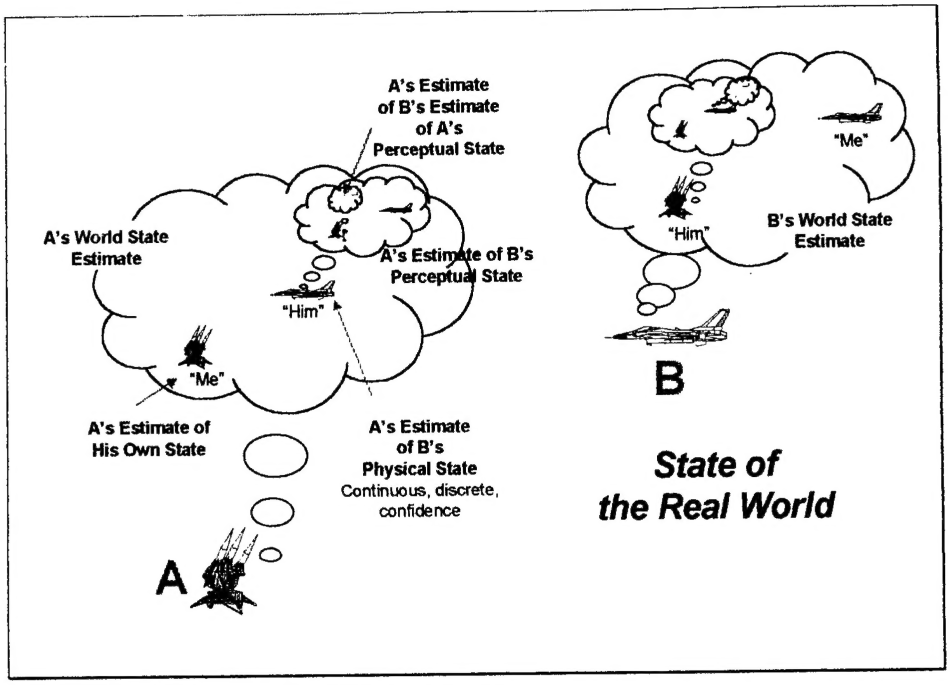

Figure 12 World states and nested state estimates

If state estimation and prediction are performed by an automatic system, then that system may be said to possess physical and perceptual states; the latter containing estimates of physical, informational and perceptual states of some aspects of the world.

5. Engineering Standards

Developing a system that utilizes existing or developmental data fusion technology requires a standard method for specifying data fusion processing and control functions, interfaces, and associated data bases. The lack of common engineering standards for data fusion systems has been a major impediment to integration and re-use of available technology. This deficiency was revealed in a Correlation Technology Assessment conducted in 1995 for the Defense Airborne Reconnaissance Office (DARO). A survey of over 50 operational and developmental Intelligence Correlation systems found a general lack of standardized — or even well-documented — performance evaluation, system engineering methodologies, architecture paradigms, or multi-spectral models of targets and collection systems. In short, current developments do not lend themselves to objective evaluation, comparison or re-use.[10]

The Air Force Space Command Space Warfare Center Project Correlation was conducted in FY1995-96 to define methods to improve the tactical utilization of data provided by current data sources by providing cost-effective means for correlating information from multiple sources.

The set of *Data Fusion System Engineering Guidelines*[11] developed under that effort were developed to provide

- a standard model for representing the requirements, design and performance of data fusion systems and
- a methodology for developing multi-source data fusion systems, selecting among architecture and technique alternatives for cost-effective satisfaction of system requirements.

The engineering guidelines recommend a data fusion architecture paradigm that defines the components of data fusion systems, their interfaces, and the systems engineering process for developing them.

The Guidelines are intended to:

- support technology infusion and operations of current and developmental national systems, broadcast services and tactical data processors;
- influence the design and operation of new national systems, services and processors; and
- support the cost-effective development and implementation of correlation systems by promoting commonality and interoperability.

As such, the Project Correlation *Data Fusion System Engineering Guidelines* can serve as a basis for affordable development, acquisition, integration and operation of multi-source/multi-sensor systems. The insights obtained into the methods used in developing practical data fusion systems provide the basis for comparisons necessary for the reuse of legacy systems and for future data fusion system developments and operations, therefore promoting cost-effective solutions.[6]

The guidelines recommended that systems performing data fusion be designed according to a particular type of architecture, using the term 'architecture' in the broad sense defined by the IEEE[12]:

> *An architecture is a structure of components, their relationships and the principles and guidelines governing their design and evolution over time.*

The general requirements for an architecture are to:

- identify a focused purpose
- facilitate user understanding/ communication
- permit comparison & integration
- promote expandability, modularity, and reusability

---

[6] The starting point for developing these guidelines was the *Engineering Guidelines for Data Correlation Algorithm Characterization*[13] developed by Llinas, Hall and Bowman under a related Project Correlation effort. Because these guidelines focused on data correlation techniques *per se*, it was necessary to extend those guidelines for the other engineering functions essential to the design and development of data fusion systems.

- achieve most useful results with least cost of development
- apply to the required range of situations.

The Guidelines recommend an architecture that represents data fusion processing in terms of nodes as shown in Figure 10. When the data fusion process is partitioned into multiple processing nodes, the process is represented via a data fusion tree, illustrated in Figure 9.

The Guidelines recommend a four-phase process for developing data fusion processes within an information processing system, shown in Figure 13.

Design and development flow from overall system requirements and constraints to a specification of the role for data fusion within the system. Further partitioning results in a specification of a data fusion tree structure and corresponding nodes. Pattern analysis of the requirements for each node allows selection of appropriate techniques, based on analysis and experience of applicability in the specified conditions.

In each phase, analysis of requirements leads to a further functional partitioning. Performance analysis of the resulting point design can lead to further analysis, repartitioning and redesign, or to initiation of the next design phase. Thus, this process is amenable to implementation via waterfall, spiral or other development methods.

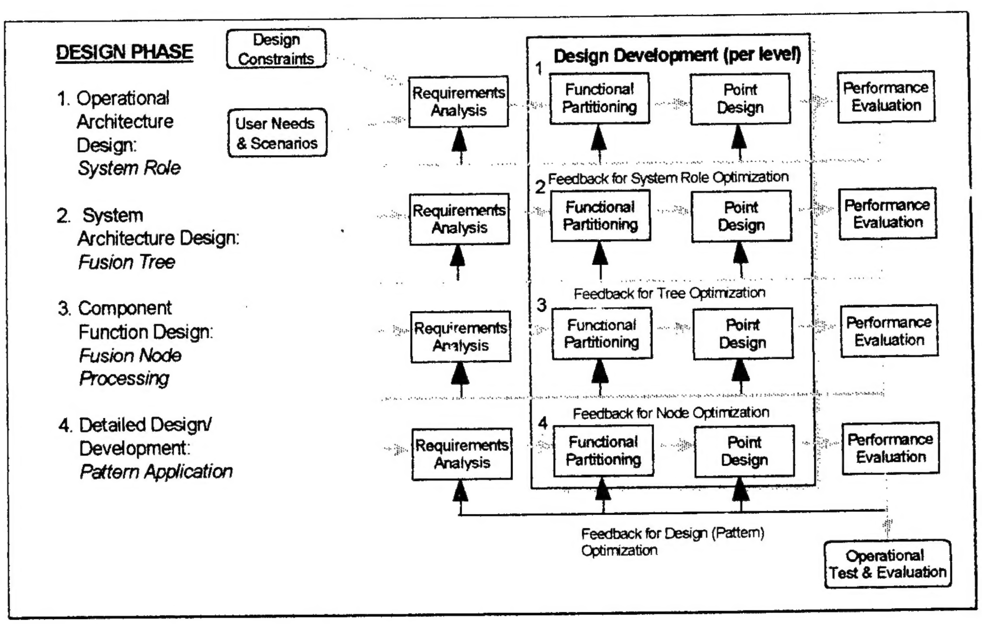

Figure 13 Datas fusion system engineering method

6. REFERENCES

[1] Franklin E. White, Jr., *Data Fusion Lexicon*, Joint Directors of Laboratories, Technical Panel for C^3, Data Fusion Sub-Panel, Naval Ocean Systems Center, San Diego, 1987.

[2] Ronald Mahler, "A Unified Foundation for Data Fusion," *Proceedings of the 1994 Data Fusion Systems Conference*, Oct. 1994.

[3] I.R. Goodman, H.T. Nguyen, and R. Mahler, *New Mathematical Tools for Data Fusion*, Artech House, 1997

[4] Shozo Mori, "Random sets in data fusion: multi-object state-estimation as a foundation of data fusion theory, " *Proc. Workshop on Applications and Theory of Random Sets,* Springer-Verlag, 1997.

[5] Lawrence Stone, Michael Finn, and Carl Barlow, *Unified Data Fusion*, Report to Office of Naval Research, Contract N00014-95-C-0052, Metron, Inc., 26 January 1996.

[6] Franklin E. White, Jr., "A model for data fusion," *Proc. 1st National Symposium on Sensor Fusion*, vol. 2, 1988.

[7] Alan N. Steinberg and Robert B. Washburn, "Multi-level fusion for Warbreaker intelligence correlation," *Proc. 8th National Symposium on Sensor Fusion* , 1995.

[8] Christopher L. Bowman, "The data fusion tree paradigm and its dual" *Proc. 7th National Symposium on Sensor Fusion*, 1994.

[9] Edward Waltz, Information Warfare: Principles and Operations, Artech House, 1998.

[10] Alan N. Steinberg, "Intelligence Correlation, 'Doing more with less'," *Proc. TMSA Surveillance & Reconnaissance Conference*, Washington, D. C. (18-19 June 1996).

[11] *Data Fusion System Engineering Guidelines*, SWC Talon-Command Operations Support Technical Report 96-11/4, vol. 2, 1997.

[12] *C4ISR Architecture Framework, Version 1.0*, C4ISR ITF Integrated Architecture Panel, CISA-0000-104-96, 7 June 1996.

[13] *Engineering Guidelines for Data Correlation Algorithm Characterization*, TENCAP SEDI Contractor Report, SEDI-96-00233, 1997.

[14] Alan N. Steinberg and Christopher L. Bowman, "Development and application of data fusion engineering guidelines," *Proc. Tenth National Symposium on Sensor Fusion*, 1997